

NASA CR- 143704

ECO 75:C-3-II

IMPACT OF REMOTE SENSING UPON THE PLANNING, MANAGEMENT AND DEVELOPMENT OF WATER RESOURCES

-Summary of Computers and Computer Growth Trends
for Hydrologic Modeling and the Impact of ERTS
Image Data Processing Load.

Peter A. Castruccio and Harry L. Loats, Jr.
ECOsystems International, Inc.
P O Box 225
Gambrills, MD 21054

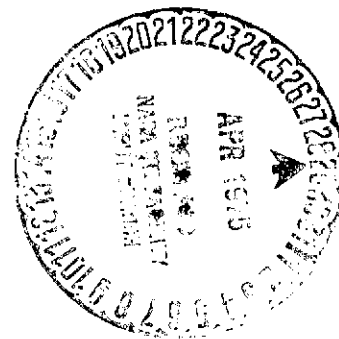
(NASA-CR-143704) - IMPACT OF REMOTE SENSING
UPON THE PLANNING, MANAGEMENT AND
DEVELOPMENT OF WATER RESOURCES. SUMMARY OF
COMPUTERS AND COMPUTER GROWTH TRENDS FOR
HYDROLOGIC MODELING AND THE (Ecosystems

N75-20802

Unclas
G3/43 18056

February, 1975
Second Quarterly Report
Contract No. NAS5-20567

Prepared for:
GODDARD SPACE FLIGHT CENTER
Greenbelt, MD. 20771



TECHNICAL REPORT STANDARD TITLE PAGE

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.	
4. Title and Subtitle Impact of Remote Sensing Upon The Planning, Management & Development of Water Resources - Summary of Computers & Computer Growth Trends for Hydrologic Modeling & the*		5. Report Date	
		6. Performing Organization Code	
7. Author(s) Peter A. Castruccio, Harry L. Loats, Jr.		8. Performing Organization Report No. 75: C-3-II	
9. Performing Organization Name and Address ECOSYSTEMS INTERNATIONAL, INC. Post Office Box 225 Gambrills, Maryland 21054		10. Work Unit No.	
		11. Contract or Grant No. NAS5-20567	
12. Sponsoring Agency Name and Address		13. Type of Report and Period Covered Type II - 2nd Quarterly Progress Report Sept. - Dec. 1974	
		14. Sponsoring Agency Code	
15. Supplementary Notes *Input of ERTS Image Data Processing Load.			
16. Abstract An analysis of current computer usage by major water resources users was made to determine the trends of usage and costs for the principal hydrologic users/models. The laws and empirical relationships governing the growth of the data processing loads were described and applied to project the future data loads. Data loads for ERTS CCT image processing computed and projected through the 1985 era. The analysis shows significant impact due to the utilization and processing of ERTS CCT's data. The effort in the remainder of the contract will center upon the assessment of this impact and the mechanisms required to optimally utilize new remote sensing data.			
17. Key Words (Selected by Author(s)) Remote Sensing, Hydrologic Computer Modeling, Water Resources Management, ERTS Image Processing Data Load		18. Distribution Statement	
19. Security Classif. (of this report) None	20. Security Classif. (of this page)	21. No. of Pages	22. Price*

*For sale by the Clearinghouse for Federal Scientific and Technical Information, Springfield, Virginia 22151.

Figure 2. Technical Report Standard Title Page

TABLE OF CONTENTS

	<u>Page</u>
Section 1: Preface	1
Section 2: Introduction	3
Section 3: Principles of Sizing Digital Computer Speed	5
Section 4: Growth Trends in Computing Power	11
Section 5: Trends of Data Processing Costs	17
Section 6: Data Processing Load and Growth Trends for Processing Hydrologic Models	28
Section 7: Data Processing Load for Processing ERTS Imagery	33
Section 8: Impact of the Remote Sensing Data Stream Upon Hydrologic Users	43

TABLES

Table 1	<u>Typical Program Mix for Measuring and Comparing Computer Power (Gobson Mix)</u>
Table 2	<u>Profile of Users of Computers for Hydrologic Purposes</u>

FIGURES

- Figure 1 Evolution of U.S. Scientific Top-of-the-Line Computers
- Figure 2 Evolution of USSR Scientific Top-of-the-Line Computers
- Figure 3 Evolution of USSR Scientific Computers
- Figure 4 Hardware/Software Cost Trends - Commercial Computers
- Figure 5 Computer Cost Composition
- Figure 6 Predicted Growth of Minicomputers 1970/1980
- Figure 7 Computer Operations Cost Trend
- Figure 8 Evolution of Computer Requirements for Hydrologic Models
- Figure 9 Evolution of Computer Requirements for Hydrologic Models
- Figure 10 Watershed Area Distribution - State and Local Users
- Figure 11 ERTS CCT-Image Processing Costs/1,000 Ha. with Pre-processing
- Figure 12 ERTS CCT-Image Processing Cost/1,000 Ha. with Pre-processor
- Figure 13 Growth of Post-ERTS Remotely Sensed CCT Imagery Processing Cost/1,000 ha.
- Figure 14 Trend of Data Processing Costs for Hydrologic Model Runs

Section 1

PREFACE

This is the second quarterly progress report of contract NAS5-20567. It analyzes the impact of remotely sensed data upon the data processing loads of hydrologic users.

The unit costs of data processing decrease with time following empirically validated laws identified herein. The trend of increasing complexities of hydrologic models grows in such a fashion as to offset the decreasing processing costs. Thus, the costs of processing hydrologic models, in the absence of new data streams, remains approximately constant with time.

The major impact of remotely sensed data upon hydrologic computing load is caused by the requirement for processing computer compatible tapes (CCT's) to extract the requisite hydrologic information.

Currently, the cost of CCT processing for typical watersheds is of the same order of magnitude as the operational costs of running hydrologic models. Because of the anticipated growth of sophistication of remotely sensed data, the CCT processing costs will also remain approximately constant with time in spite of the historical decrease of computing costs.

The logical consequence of these trends will manifest itself as follows:

1. The small users will lag the larger users in incorporating remotely sensed data due to the cost of processing CCT's.
2. The small users will rely more heavily upon direct photo-

interpretation techniques, with consequent reduced effectiveness caused by the loss of radiometric information.

3. Only the very large users with ready access to sophisticated data processing facilities will be able to immediately exploit remote sensing to its full potential.

Because of the large population of small users, NASA should consider remedial measures to alleviate the small user-large user information gap.

The next quarterly progress report will set forth guidelines and recommendations to this effect.

Section 2

INTRODUCTION

Extensive research in the ERTS program has indicated that the area of water resources is potentially very valuable. The utility of remote sensing data for both hydrologic planning and management models has been demonstrated; the effort to optimally use remote sensing information is continuing with LANDSAT.

The value of remote sensing from space is evidenced by the recent launch of the Soviet earth resources satellite PRIRODA and by the intensive studies by foreign nations: ESA (ESRO), France, West Germany, Canada and Japan.

NASA is planning to implement additional earth resources satellites, including increased capability to perform hydrologic observations.

International interest can be gauged by the significant growth of ground facilities for the reception, analysis and dissemination of earth resources satellite data. Brazil, Canada, Italy, and the Federal Republic of Germany have purchased ground facilities. The USSR is reported to have implemented its own ground system. Norway, France, Iran and the United Nations are in the advanced design stages.

In the first quarterly progress report the hydrologic users were identified and their uses, data processing equipment and models were analyzed. This second quarterly progress report treats the impact of the remotely sensed data stream upon the user data processing facilities. Specifically: (1) the growth trends in computing power;

(2) data processing cost trends; (3) current and future hydrologic modeling data processing requirements; (4) ERTS remotely sensed data processing requirements and growth trends.

Section 3

PRINCIPLES OF SIZING DIGITAL COMPUTER SPEED

Computing power is commonly defined in two ways: (1) Internal Performance, which is the computing speed of the Central Processing Unit (CPU). This definition tacitly assumes that the I/O is of infinite capacity; (2) Throughput, which is the speed of the system, including CPU and Input-Output (I/O) peripherals. Throughput never exceeds Internal Performance.

We will here utilize primarily the definition of Internal Performance. The reason is that measurements of throughput require specification of the I/O configuration used, and of the problem being run. This information is difficult and costly to obtain, and not really needed for the "plus or minus three decibel" type of overall technological assessment that will be made here.

There is no general agreement in the trade, or at any international level, on the units of measurement of internal performance. The most used units and their corresponding methods of measurement are:

- (1) Benchmark timings, i.e. the time required to process specific, defined problems. This is by far the most accurate method, used frequently to select machines competitively, but not practical for general comparisons. The reason is that data on benchmark timings are scarce because these measurements are quite expensive.

- (2) Instructions per second (IPS, and multiple KIPS and MIPS).

One constructs a set of programs, "representative" of typ-

ical spectra of scientific problems, and measures the "average" speed with which the CPU processes them. Strictly speaking, the method is exact only when comparing machines whose characteristics are roughly similar. Otherwise, one may find that Machine B which is slower than Machine A on the "representative" program may actually perform faster on actual problems. Nevertheless, KIPS and MIPS are becoming the yardstick of performance most used in the industry.

- (3) Operations per second. Similar to (2) above in concept. In general, depending upon the type of instructions and upon the architecture of the machine, in scientific applications one operation requires more time than one instruction.* For purposes of across-the-board comparison, a reasonable average figure is: 2 Instructions per second = 1 Operation per second; 1.5 Additions per second = 1 Operation per second.

*That this is so, can be seen from a simple consideration. Take for example the operation of addition. What the instructions must do is to cause the machine to fetch both addends from memory, then add them together, and finally to return the result to memory. In a single-address machine, for example, this requires typically a LOAD instruction, then an ADD, then a STORE. Three instructions per operation. In double- and triple-address machines, one instruction suffices. (For example, ADD A to B and store in C are all performed from a single three-address instruction). However, the time it takes to perform an operation of addition, or multiplication, or worse yet, division, is generally longer than the time required to perform a logical operation such as STORE or MOVE. Again, the time required depends on whether the operation must be done with single or double precision. Double generally takes longer, depending upon the design of the machine.

It should be noted that this method of comparison is too coarse for precisely judging the relative performance of two machines for purposes of deciding which one to acquire. However, when applied to the charting of secular technological trends, experience has shown that the method works quite well, provided that a sufficient variety of machine models is included in the comparison. The reason is that errors in assessing individual machine performance tend to cancel out statistically over the large population of machine models.

The ways in which the speed of a machine is measured or estimated is:

- a) To actually measure the time required to run a specific program. This is known as benchmark timing.
- b) To test the machine against typical mixes of programs. Widely used is the Gibson Mix, whose composition is shown in Table 1.
- c) To break the program into its individual instructions in BAL (Basic Assembly Language); calculate the mix of instructions; go back to the machine specification sheet & determine the speed of each instruction; and finally determine the total speed. This is a rather laborious procedure if there are many different programs to be considered.
- d) To take an average, based on the general characteristics of the program. For avionics and aerospace programs of guidance and fire control, a widely employed measure of speed is to take the average between 4 additions and one multiplication time and divide the total time by 5. The

Table 1
TYPICAL PROGRAM MIX FOR MEASURING AND COMPARING
COMPUTER POWER (GIBSON MIX)

<u>FUNCTION</u>	<u>WEIGHTING</u>
Fixed Point	
Add/Subtract	0.330
Multiply	0.006
Divide	0.002
Branch	0.065
Compare	0.040
Transfer 8 Characters	0.175
Shift	0.046
Logical	0.017
Modification	0.190
Floating Point	
Multiply	0.040
Add	0.073
Divide	0.016

result is taken to be the time required per operation.

This yields the speed of the machine, not in kips, but in a somewhat different measure, known as kops (operations per second rather than instructions per second).

The above definitions of "internal performance" are applicable for programs in which there is a lot of internal number manipulation, with little input/output load. If the I/O load is significant, the correct measure is that of "throughput," which is always smaller than "internal performance." The degradation between internal performance and throughput depends upon whether the input rate or the output rate exceeds the machine's internal performance. In most cases, the bottleneck arises from output rate.

An idea of why this happens can be had as follows: assume that the program requires a lot of printing. Assume a typical high-speed printer of 1,500 lines/minute (25 lines/second). This means that every time the machine is required to print a line (regardless of how full the line is), it consumes 1/25th of a second. For a hundred kip machine, this is equivalent to consuming a time lapse of 100,000 divided by 25 or 4,000 instructions.

It is clear that if the machine must continuously print, no matter how fast it is internally, the throughput cannot exceed the number of instructions required to generate one line.

The throughput in this case is calculable from knowledge of the printout format.

Likewise for the input: conventional magnetic tape can feed approximately 125,000 bytes/second. If each byte calls for n instructions, the machine is required to perform 125,000 times n ips.

If the internal performance is slower than this, the machine will slow down.

For programs written in Fortran, a widely used assumption is that one Fortran statement is equivalent to between four and 10 BAL (Basic Assembly Language) instructions. This assumption suffers from the same inaccuracies discussed above. For example, DO loops may require tens and up to hundreds of instructions. To achieve greater precision, one should count the number of Fortran statements in the program and the corresponding numbers of BAL instructions pertaining to each statement.

Section 4

GROWTH TRENDS IN COMPUTING POWER

The principal criteria of merit of data-processing systems are:

- (1) Computing power - - the speed at which the system performs computations.
- (2) Reliability or "up-time" - - the productivity ratio of the system: hours worked divided by total hours available.
- (3) Memory size - - the maximum available memory.
- (4) Price/performance - - the price of the data processing installation, divided by its computing power. This has been shown to have a definite relationship to machine power and year of entry into the market (Grosche's Law).
- (5) Software complement - - number and quality of programs supplied with the machine.
- (6) Compatibility - - the ease with which the software can be applied to other models of the same manufacturer's line, or generally available on the market.
- (7) Growth - - what next larger model is or will be available.
- (8) Technology - - the type of circuits employed. This is an indicator of "modernity."

When attempting to forecast evolution, the most comprehensive indicator is computing power. The reason is simple. A high-power computer is only practical if: (1) its size is not unreasonably large, implying the existence of a technology of "reasonable" compactness. For example, a 360/75 could never be built out of vacuum

tubes; (2) its reliability is tolerable (implying a technology of sufficiently high circuit reliability so that ensembles of 50,000 to 100,000 circuits, typical of large machines, are still reasonably proficient); (3) memory size is at least minimally adequate for the problems the computer is designed to solve (too small a memory would reduce the computing power of the machine, thus rendering its development somewhat pointless); (4) the price is reasonable.

In conclusion, the indicator "computing power" contains much implicit information regarding the other indicators: Technology, Reliability, Memory Size, Price.

Figure 1 plots the internal performance, in operations per second, of the U.S. top-of-the-line general-purpose scientific machines as a function of the year of first installation.

The top-of-the-line is the set of the most powerful machines. It is indicative of the "best" hardware that it is practical to produce in any one era. In the U.S., under the stimulus of demand and of improving technology, the growth of the top-of-the-line, independently of manufacturer, has followed over the last 20 years the empirical relationship:

$$P = P_1 \times 2^{0.5 \times (t-t_1)}$$

or

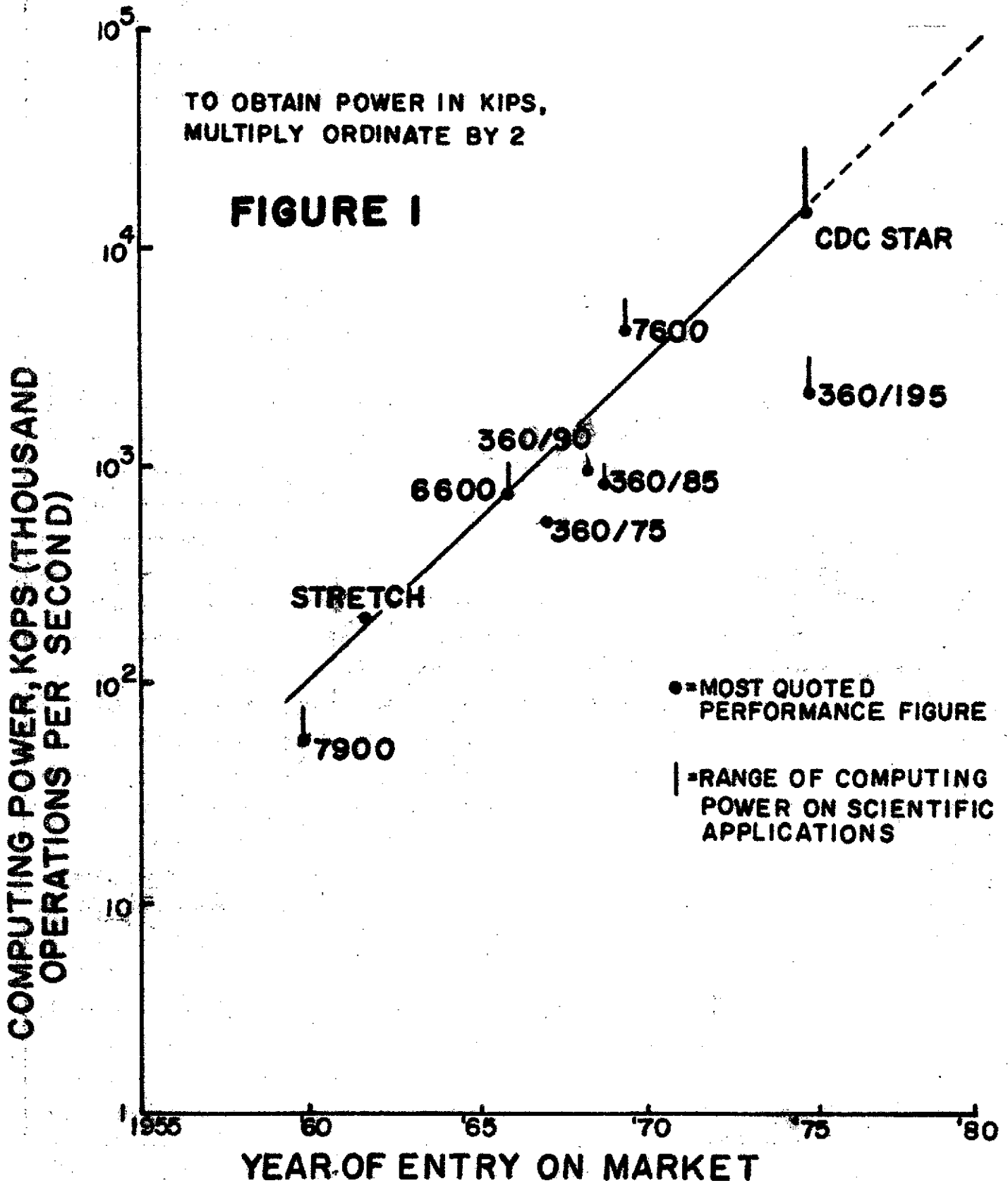
$$P/P_1 = (\sqrt{2})^{(t-t_1)}$$

where: P = computing power in year t

P_1 = computing power in year t_1

This says in essence that technological progress has grown at such a pace that the power of the fastest computers has doubled every

EVOLUTION OF US SCIENTIFIC TOP-OF-THE-LINE COMPUTERS



two years.

It should be noted that this is a secular trend: it does not predict exactly when a specific growth machine will see the light, nor does it pinpoint the exact computing power of the most powerful machines within a given time frame.

As an interesting comparison, the USSR (the next major producer of big machines after the U.S.) trend is plotted in Figure 2 and compared with the U.S. Note that the slopes, i.e. the growth exponents, are approximately the same for the U.S. and USSR.

As we shall see in the next section, the cost of processing is least when the top-of-the-line (hereinafter referred to as TOL) is employed. Thus the computing power of the TOL is also an excellent indicator of data processing costs.

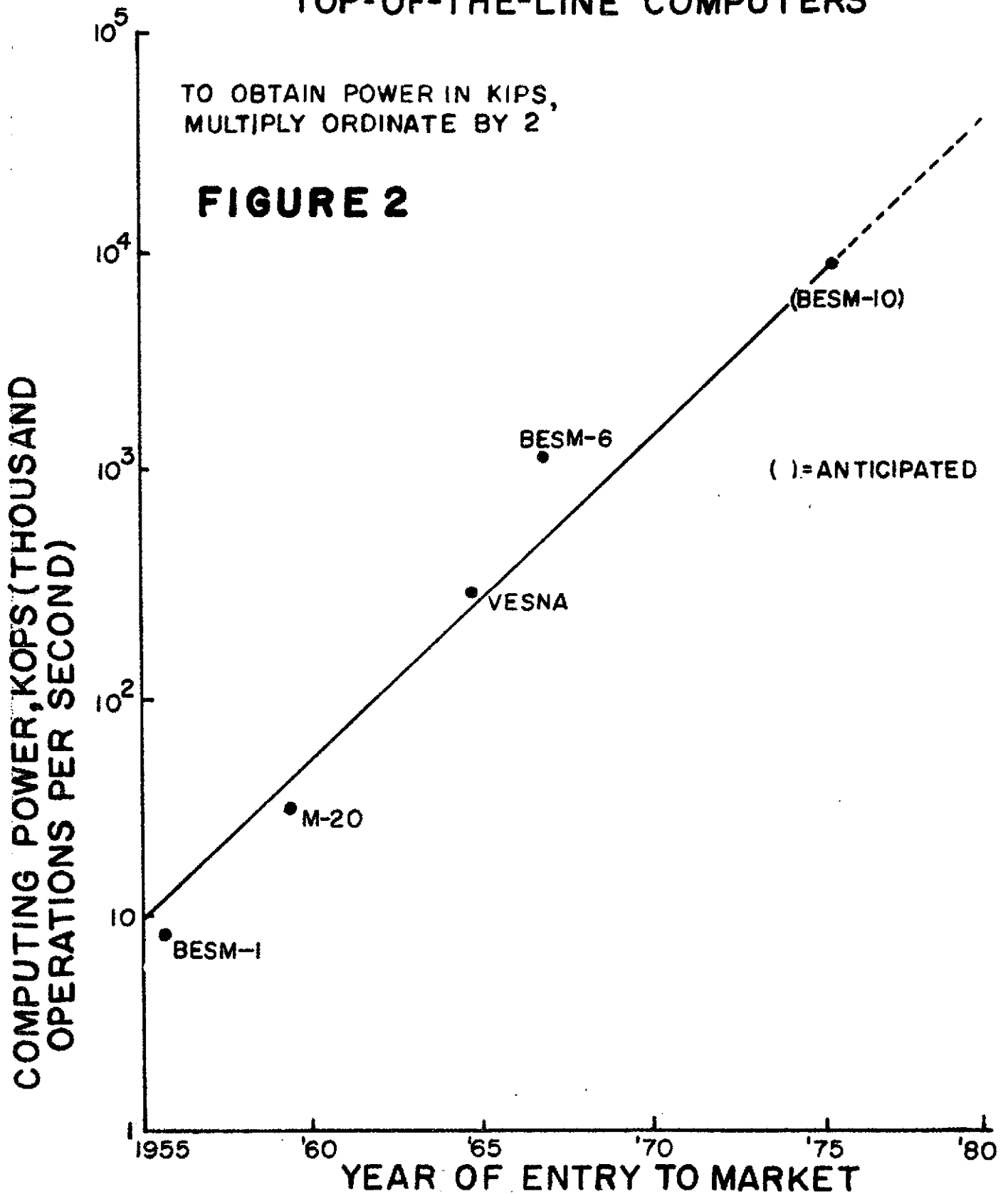
Of course, manufacturers do not confine themselves to producing the TOL class of machines. The region below the TOL is populated at any one time by several machines of lesser power, which span the gap between the TOL and the minicomputer class.

The U.S. machine population is well known. Of some interest is Figure 3, which depicts the USSR population of machines below the TOL.

EVOLUTION OF USSR SCIENTIFIC TOP-OF-THE-LINE COMPUTERS

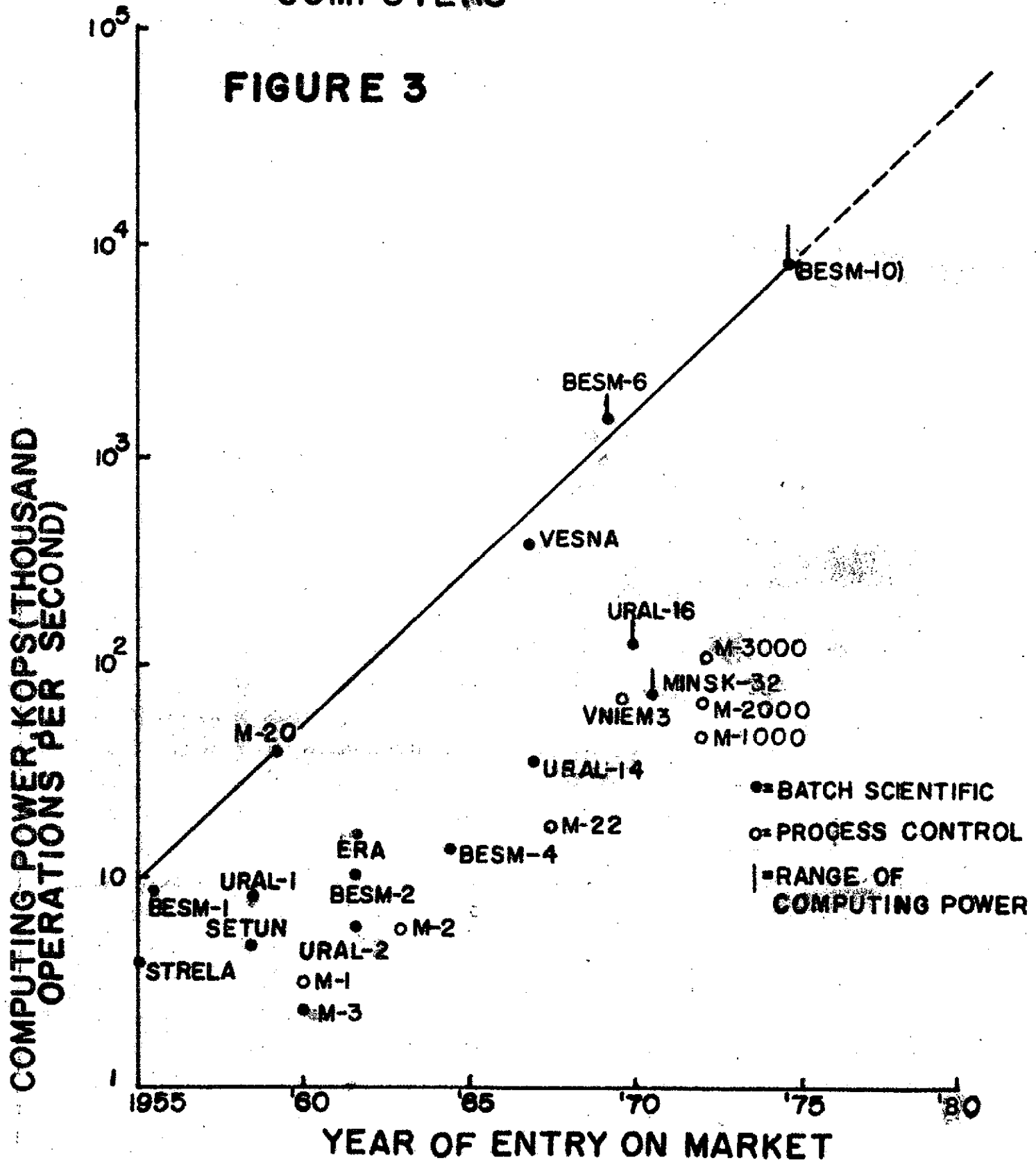
TO OBTAIN POWER IN KIPS,
MULTIPLY ORDINATE BY 2

FIGURE 2



EVOLUTION OF USSR SCIENTIFIC COMPUTERS

FIGURE 3



Section 5

TRENDS OF DATA PROCESSING COSTS

A universally used measure of the economic effectiveness of data processing equipment is price-performance, defined as the cost per instruction executed, or equivalently, the number of instructions executed per dollar.

The principal trends of interest in the evolution of computer economics are:

- 1) Grosche's law, which should more properly be referred to as Grosche's empirical relationship. It holds that, on the average and at any moment in time, the price of a computing machine is proportional to the square root of its computing power. This means that a high-priced machine performs more instructions per dollar than a smaller, lower-priced machine. As a typical example, the 360/195 complete system did cost at its point of entry to the market (1970) typically and approximately \$10 million. Its average speed is 6 MIPS. The 360/65 system did cost in the same year \$3 million. Its average speed is 0.65 MIPS. It can be seen that the ratio of speeds, $\frac{6}{0.65} = 9$ is approximately the square of the prices: $\frac{10}{3} = 3.3$. This relationship has held approximately true since the early 1950's. This means that the price-performance is better (more instructions per dollar, or less dollars per instruction) for large than for small machines. The obvious question is: why doesn't everybody use large machines? The answer is equally obvious: because they cannot afford

the investment. As a matter of fact, some of the large users employ large machines for hydrologic processing, sharing this application with many others. The small user does not have that many other applications to warrant acquisition of large computers.

- 2) The law of the top-of-the-line, which again is not a law, but a historical trend which has held since the early 1950's. It states that the top-of-the-line (i.e. the largest machine which enters the market) increases in power by $\sqrt{2}$ every year. In other words, computer power doubles every second year. The growth of the smaller machines is somewhat slower, of order $\sqrt[3]{2}$ per year approximately.
- 3) The combination of these two relationships indicates that the cost of the top-of-the-line remains constant. In fact, since the early nineteen-fifties, the cost of the most powerful machine purchasable at any one time has remained at the approximate level of \$10 million.
- 4) The historical cost decrease. On the average, the price for equal computing power (MIPS) decreases by a factor of 0.75 every year.

$$P = P_0 (0.75)^{(t-t_0)}$$

where:

P_0 = price in year t_0

P = price in year t

Combining Grosche's law 1) with the historical cost decrease 4) shows that the price-performance with time of any machine can approximately be expressed as:

$$C = C_0 \sqrt{\frac{P_0}{P}} (0.75)^t$$

where:

C = price of machine of power P at future
time t

C₀ = price of the TOL machine at time t₀

P₀ = power of TOL machine at time t₀

Note that the above are simply historical trends, which have been observed in retrospect over the last 25 years. Nothing guarantees that they will hold in the future: recent trends indicate some departure from these "laws," in the growth of the TOL. For example, extrapolation of the TOL trend to 1975 indicates that there should appear, this year, a commercial machine capable of approximately 120 MIPS. No such computer is available. To be sure, IBM was planning a 100-MIP machine for this time frame: this was eventually discontinued. ARPA was at one time planning a 200-MIP plus version of the ILLIAC IV, due approximately 1976 or 1977. The effort has been slowed down. The reason why the TOL trend is slowing down is that TOL machines, since the days of ENIAC, have been motivated by the Government market: weather forecasting, nuclear effects, ballistic missile defense, and similar requirements. Commercial requirements are primarily in seismic exploration. Under present conditions, the market is small and aleatory. Thus, commercial manufacturers prefer to invest their resources in the smaller and more saleable machines. It is difficult to foretell whether the TOL trend will

change in the next several years. However, the growth of the second-and-third echelon machines below the TOL still appears to follow the "doubling-every-two-years" trend.

It should further be noted that these trends hold only when averaged over the entire U.S. market. They do not imply that any one manufacturer will automatically enter the market, year after year, with machines exactly obeying the general trend. In fact, individual manufacturers tend to produce "generations" of machines, which remain constant over several years. Competition between manufacturers causes the various generations to interleave in time. Various other economic trends have been observed, more general and softer than those previously reported.

- 5) The migration trend, which can also be stated as an aspect of Parkinson's "law:" work expands to fill the computer, or stated more pessimistically "computers never save money." What this means is that, even though a computer is often purchased for the specific objective of saving labor in a defined operation, such as payroll or modeling, its availability unavoidably causes the user to try things never tried before. Thus, the original intended use expands into evermore sophisticated uses not contemplated at the time of purchase. The ever-expanding requirements, coupled with the historical reduction of price, motivate the user to periodically acquire a more

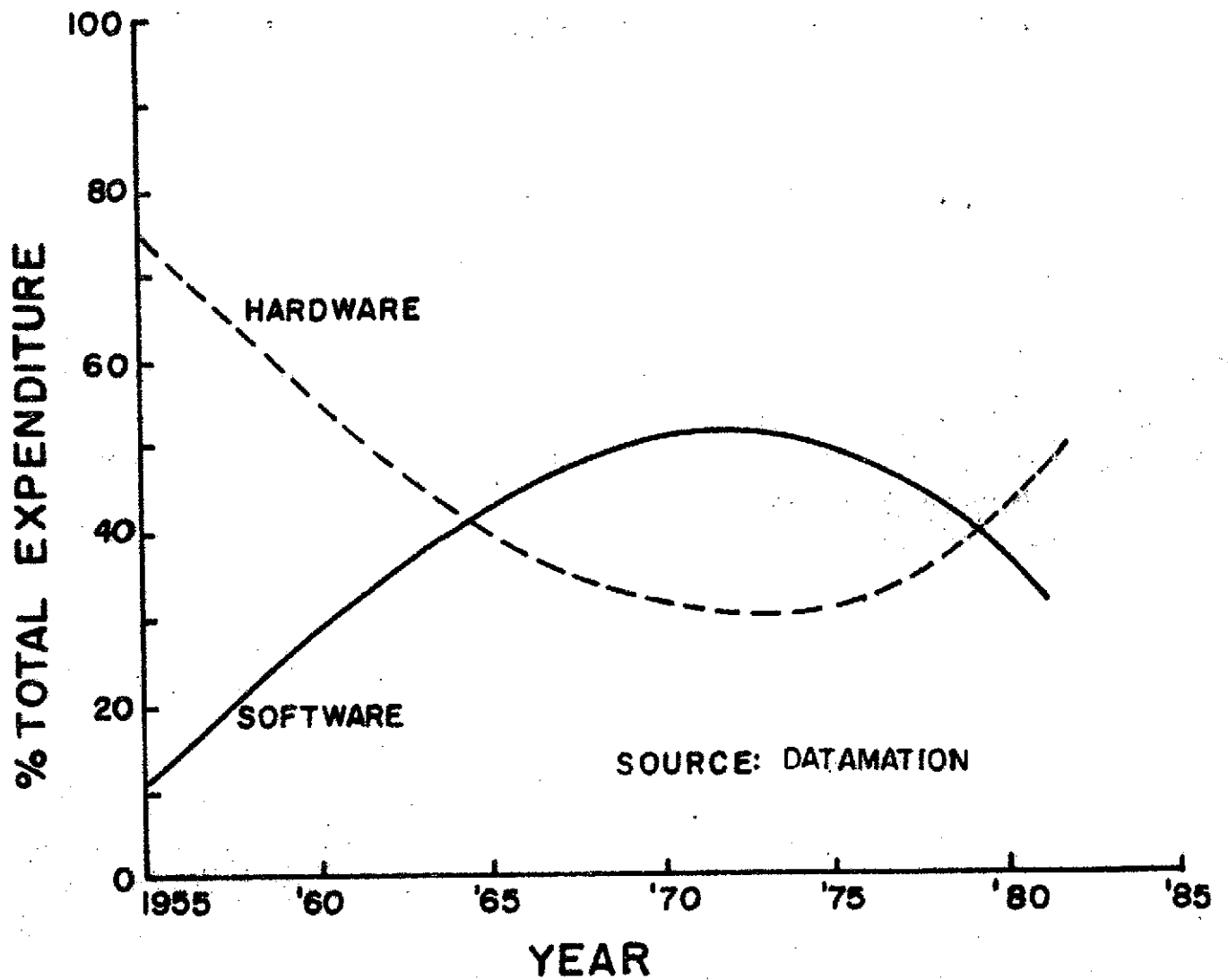
powerful machine. Thus, the user's computing power tends to "migrate" upwards. At the same time, the complexity of the application also migrates upwards.

The consequences of this trend, specifically in hydrologic modeling, and possibly in image data processing, are that models and processing algorithms tend to grow apace with the expanding power of the machines. This trend is charted in the next section.

- 6) The size of fast available memory, for a given price level, grows with computing power. No hard and fast rules exactly quantify this growth, particularly since many users do not employ the maximum available memory for a given machine. A gross relationship is that the largest available memory grows as the cube root of computer power.
- 7) Hardware-software mix. In the early fifties, hardware costs represented the major share of data processing costs. Since then, the combination of decreasing hardware costs and increasing programming sophistication and programmer wages have shifted the hardware-software mix towards the fifty-fifty point. Forecasts for the future vary: for large, complex systems, by 1980 the software is expected to constitute 80% of the data processing costs. More significant for hydrologic applications is the forecast for all systems shown in Figure 4.

FIGURE 4

HARDWARE/SOFTWARE COST TRENDS - COMMERCIAL COMPUTERS



This indicates a renewed climb of the hardware costs, mostly due to the expansion of peripheral equipment.

- 8) The trend towards increased peripherals is depicted in Figure 5. It is induced by increasing emphasis upon interactive systems, increased use of computers as communications switching and input-output devices, use of large buffer memories, and expanding employment of time-shared systems.
- 9) The decreasing hardware costs have prompted the increase of minicomputers, wherein the term mini is strictly relative to the larger machines. The power of many minis current is equal to or larger than that of the top-of-the-line of the mid-fifties. The cost trend for minis is shown in Figure 6 .

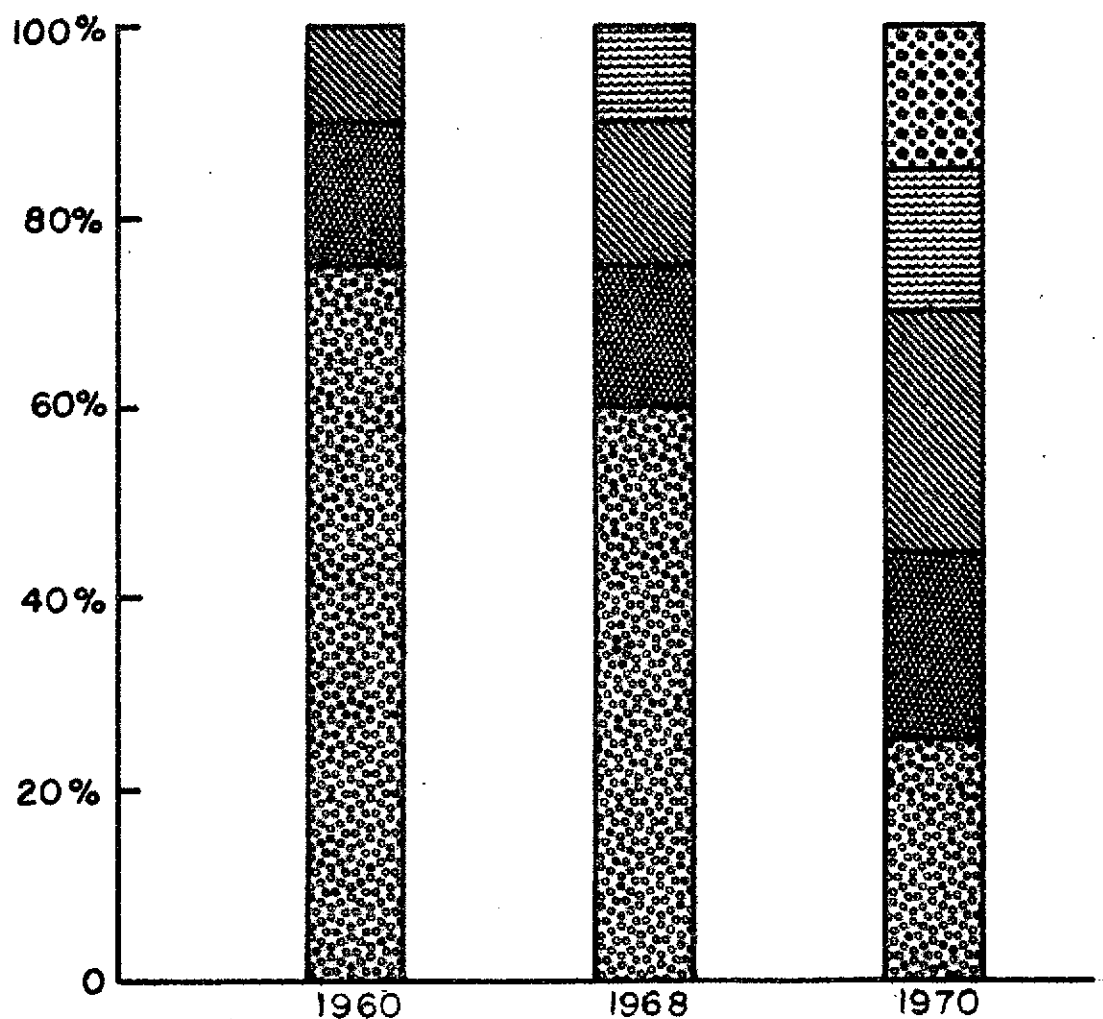
Note the large increase forecasted for data logging, switching and acquisition functions, and for process control (real-time) functions. The increase forecasted for scientific applications such as hydrologic modeling is, however, modest.


Figure 7 synthesizes the historical trend of computing costs. The parallel straight-line boundaries in the figure indicate the range of computing power, which has been employed for the more sophisticated hydrologic models (mostly processed on a shared basis). Items 11 and 12 in Figure 7 are small computers, which have been used in simpler hydrologic models. In particular, the IBM 1130 has found relatively wide application for river forecasting in the ESSA (now NOAA) organization.

FIGURE 5

COMPUTER COST COMPOSITION

SOURCE: OECD



 = CPU

 = Alpha-Numeric File Memory

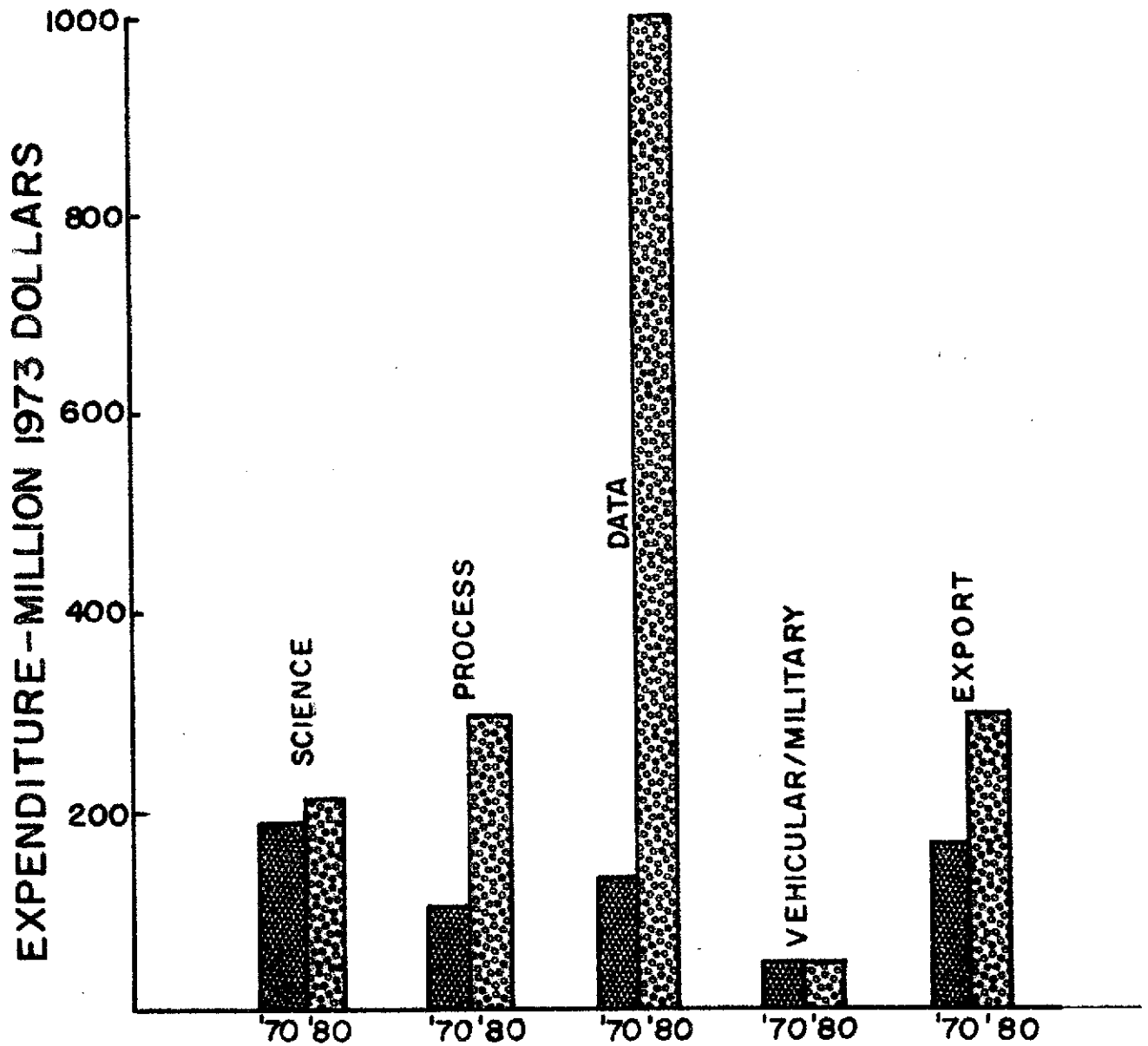
 = I-O

 = Data Transmission

 = Video File Storage/Processing

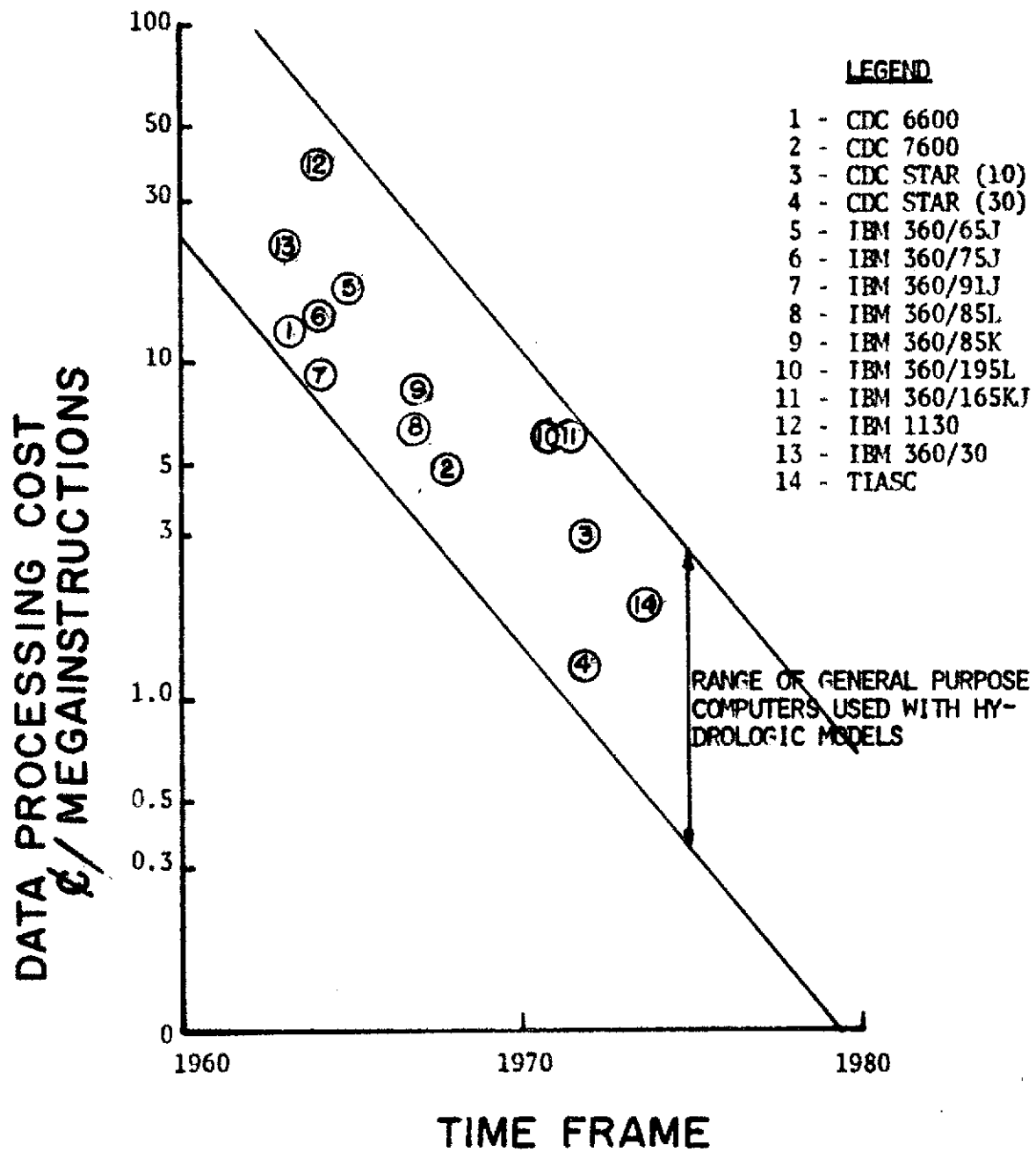
FIGURE 6

**PREDICTED GROWTH OF MINICOMPUTERS
1970/1980**



SOURCE: DATAMATION

FIGURE 7
COMPUTER OPERATION COST TREND



We can conclude that if the historical trend experienced over the last two and a half decades continues, and, barring inflationary distortions, by 1980, the cost of processing should come down to between one and five cents per million instructions.

As a final note, one must remember that Figure 7 reflects the processing costs only. To these must be added the costs of readying the data for computer usage, plus the costs of developing the software,

The costs of readying the data involves the standard functions of aerial photo interpretation, digitization of rain and streamflow records, measurement of streamlengths and other parameters of interest from maps, aerial photos or ERTS imagery.

The costs of data preparation by manual means is not estimated here. The cost of automated data interpretation from ERTS- derived computer-compatible tapes is presented in a subsequent section.

The cost of developing the software is generally high. For this reason, by and large only Federal agencies and some of the larger and wealthier States have performed this function, and will in all probability continue to do so. The intermediate and small users will continue to employ standard, already developed software. Since hydrologic models are mostly developed on U.S. Government funds, they are public property and, therefore, their cost to users is essentially nil.

In summary, the cost of processing the hydrologic models, shown in this Section, plus the costs of automated interpretation of ERTS imagery presented in a subsequent section, are good indicators of the impact of remote sensing upon hydrology users.

Section 6

DATA PROCESSING LOAD AND GROWTH TRENDS FOR PROCESSING HYDROLOGIC MODELS

The information gathered during the previous reporting period has been synthesized in Table 2 into profiles by distinct classes of users of hydrologic models. Note the ascension of the computing power available to the users: the power of the available machines increases with the size of the user. Similarly, the magnitude of the hydrologic programs grows with the size of the user. Note that the program sizes are given in terms of Fortran statements: this number must be multiplied by at least a factor of four and up to ten to obtain the program size in terms of BAL instructions.

In practice, to obtain the hydrologic behavior of a watershed, each program is run not just once, but several times, to allow for calibration, setting of constants, statistical checks, and so forth. Thus a good overall measure of the program's length is the total number of BAL instructions required to perform a complete set. This number equals the number of BAL program instructions times the number of runs, plus the overhead required to set up and calibrate. The information gathered during the previous reporting period was collated to assess the trend of growth of hydrologic models. The results are depicted in Figures 8 and 9. Note that program load grows versus time. This is not surprising: it simply confirms the trend of expanding use (a form of Parkinson's law) indicated in the previous Section. The last point on the curve of the Figures is an estimation of the program load required by evolution of current hydrological programs towards the direction of microhydrology.

ORIGINAL PAGE IS
OF POOR QUALITY

PROFILE OF USERS OF COMPUTERS FOR HYDROLOGIC PURPOSES							
DESCRIPTION	SCOPE	PERSONNEL COMPLEMENT	COMPUTER COMPLEMENT	PROGRAM SIZE (NUMBER OF FORTRAN STMTS)	POTENTIAL DIRECT R/S INPUT	POTENTIAL DIRECT DCP DATA INPUT	COMMENTS
1 - <u>SMALL GENERALIST</u> ◦ CIVIL ENGR. CONSULTANTS EX: MR. T. BALLARD	◦ SMALL SCALE HYDRAULIC STRUCTURES DESIGN ◦ SOIL SURVEYS ◦ DRAINAGE DESIGN	◦ ENGINEER WITH LIMITED PROGRAMMING CAPABILITY	◦ TERMINAL TO TIME SHAPE SMALL - MED. COMPUTERS ◦ MINICOMPUTERS (PDP 12)	$10^2 - 10^3$	_____	_____	DATA COLLECTED FROM LARGER CATEGORIES: WILL EVENTUALLY FILTRATE THIS LEVEL.
2 - <u>SMALL SPECIALISTS</u> LG. CIVIL ENGINEERS EX: WHITHAM REQUARDT, PARSONS BRINCKERHOFF, UNIVERSITIES	◦ HIGHWAY PROJECTS ◦ URBAN/REGIONAL PUBLIC WORKS ◦ USE WATERSHED MODELS DEVELOPED BY OTHER AGENCIES	◦ ENGINEERS WITH LIMITED PROGRAMMING CAPABILITY ◦ PROGRAMMING SPECIALISTS IN LARGER ORGANIZATIONS ◦ HYDROLOGY SECTIONS	◦ TERMINALS TO MED. TIME SHARE ◦ MINI - MED. COM- PUTERS (IBM 2130)	$10^2 - 10^3$	◦ ANALYSIS OF PARTIAL ERTS FRAMES PICTORIAL FOR HYDROLOGIC LAND USE INFORMATION	_____	WILL POTENTIALLY BE A MARKET FOR UP-TO-DATE- INFORMATION ON SMALL HY- DROLOGIC AREAL UNITS.
3 - <u>LARGE LOCAL SPECIAL- ISTS</u> ◦ POWER COS. ◦ HYDROCOMP ◦ WRG ◦ WATER RES. INST. STATES	◦ REGIONAL/SUB REGIONAL MODELING ◦ ENVIRONMENTAL IMPACT RESPONSE ◦ PARAMETRIC MODELING OF WATERSHED RESPONSE	◦ PROGRAMMING SPECIALISTS ◦ HYDROLOGIC DATA/MOD- ELING SPECIALISTS	◦ TIME SHARING OF MED - LG. COMPUTERS (IBM 360-20, 50) ◦ IN-HOUSE COMPUTERS COMPLEMENT MULTI- MED ◦ ANALOG DIGITAL DATA COLLECTION NET	$10^2 - 10^3$	◦ NEW DATA STREAMS FOR PLANNING PARAMETRIC MODELS (EXISTING) ◦ INPUTS FOR NEW PLAN- NING PARAMETRIC MODELS ◦ WATERSHED DATA UPDATES FOR HYDROLOGIC LAND- USE DATA	_____	LARGEST IMPACT ON PLAN- NING AND MANAGEMENT OPS
4 - <u>LARGE REGIONAL SPEC- IALISTS</u> ◦ MID ATLANTIC REG., COE ◦ TVA ◦ USGS REG. OFFICE ◦ STATE WATER RES. AFFICIES ◦ REG. POWER ADMIN.	◦ LARGE PUBLIC WORK DESIGN/CONSTRUCTIONS ◦ COMPIATION OF REGIONAL HYDROLOGIC STATISTICS ◦ REAL TIME DATA COLLEC- TION/ ◦ REAL TIME OPERATIONS	◦ SAME AS ABOVE	◦ LINKED TO LG. SCALE COMPUTER NETS ◦ IN-HOUSE COMPUTERS- MED ◦ ANALOG DIGITAL DATA COLLECTION/ (IBM 360, 370)	$10^2 - 10^3$	"SAME AS FOR ITEM 3" PLUS DATA FOR MANAGE- MENT/OPERATIONS	"SAME AS ITEM 3"	LARGEST IMPACT OF MODEL DEVELOPMENT
5 - <u>FEDERAL AGENCIES</u> ◦ HEC, CTE ◦ NASRCS	◦ HYDROLOGIC MODEL DEVEL- OPMENT ◦ TRAINING CENTER OPS ◦ MULTI-BASIN DEVELOP- MENT	◦ PROGRAMMING STAFF ◦ HYDROLOGIC DATA/MODEL- ING SPECIALISTS ◦ OPERATING/TRAINING CENTER	◦ MULTIPLE MED. SCALE COMPUTERS (IBM 360, 70) ◦ OPERATE COMPUTER NETS	$10^3 - 10^4$	"SAME AS ITEM 3"	LOW IMPACT GENERALLY HANDLED AT REGIONAL LEVEL	LARGEST IMPACT OF MODEL DEVELOPMENT

FIGURE 8

**EVOLUTION OF COMPUTER REQUIREMENTS
FOR HYDROLOGIC MODELS**

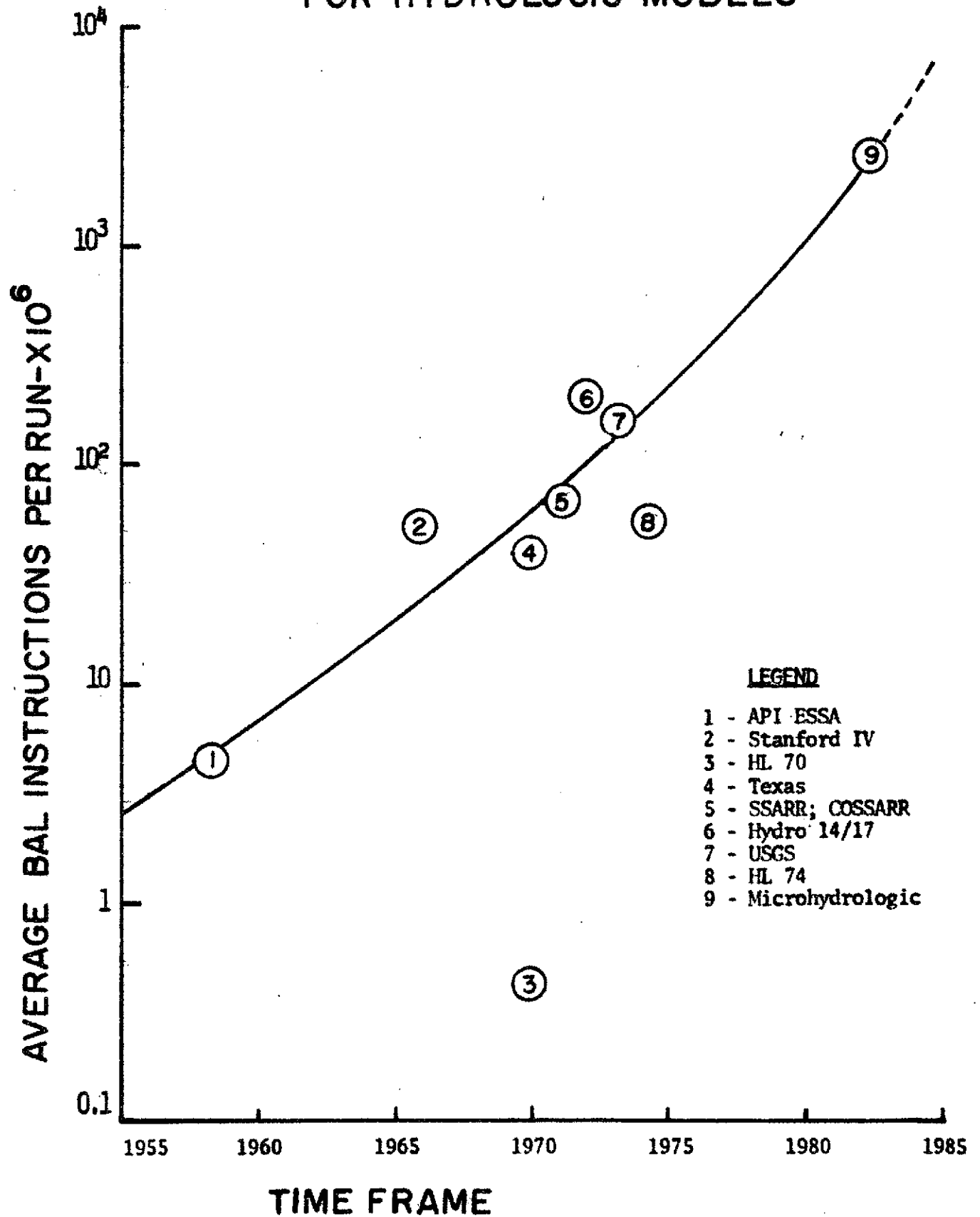
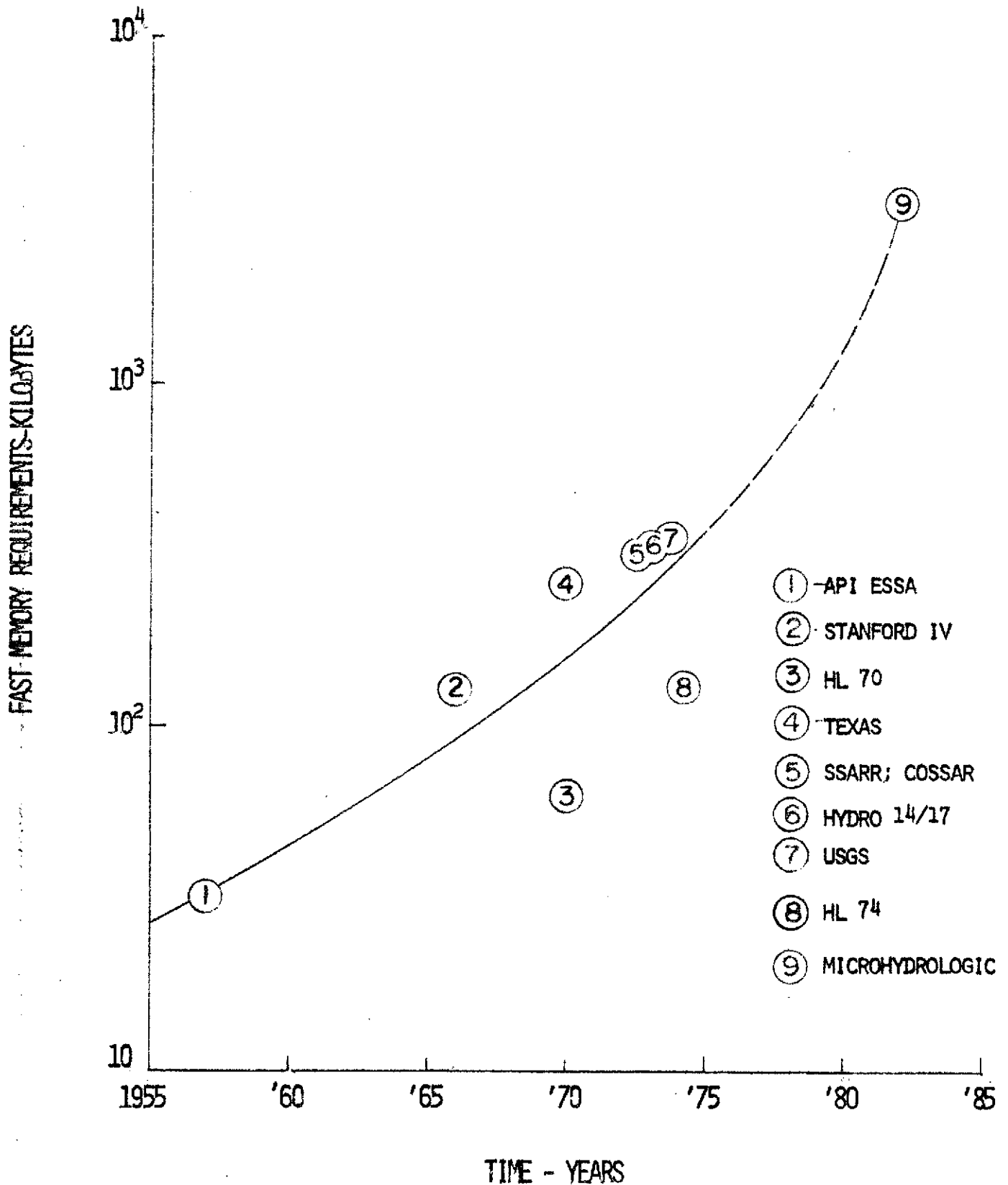


FIGURE 9

EVOLUTION OF COMPUTER REQUIREMENTS FOR HYDROLOGIC MODELS



The above evolutionary trend applies to rainfall-runoff models. The additional load imposed by advanced applications such as soil moisture accounting, will also be interesting to evaluate.

Section 7

DATA PROCESSING LOAD FOR PROCESSING ERTS IMAGERY

With present state-of-the-art algorithms, the number of instructions required to assign each pixel to a class is approximately 1,000 per band. Complete pixel-by-pixel processing of one ERTS frame (3.5 million hectares), in four bands, requires approximately $4 \times 9 \times 10^6 \times 1,000 = 36 \times 10^9$, or 36 billion instructions (since one ERTS frame contains approximately 9 million pixels). In addition, some overhead must be added for training of the computer, and for the operating system. Further overhead is required for special processing functions such as border recognition. A reasonable rule-of-thumb for the overhead required for these functions (sophisticated processing) is a factor of two.

To give a feel for these numbers, consider the time required to process an ERTS frame in four bands on a large machine, the IBM 360/75: 10 hours without overhead, 20 hours with sophisticated processing.

To completely pixel-by-pixel process an area of 1,000 hectares, simple computations show that the number of instructions required is:

Without overhead: 9 million instructions

With overhead : 18 million instructions

The equivalent 360/75 processing times required are:

Without overhead: 8 to 10 seconds

With overhead: 16 to 20 seconds

The processing time for 1,000 hectares can serve as the basis

for judging the processing time for watersheds. The area distribution of watersheds of importance to State and local users is shown in Figure 10. It indicates that the median watershed area is 10,000 hectares, ranking up to a maximum size of order 50,000 hectares. Watersheds of interest to Federal users range much higher.

Since pixel processing is a highly repetitive procedure, it lends itself to so-called vector processing, or preprocessing. A preprocessor is a hard-wired (or microprogrammed) machine, which can be configured to perform sequences of the same operation at high speeds.

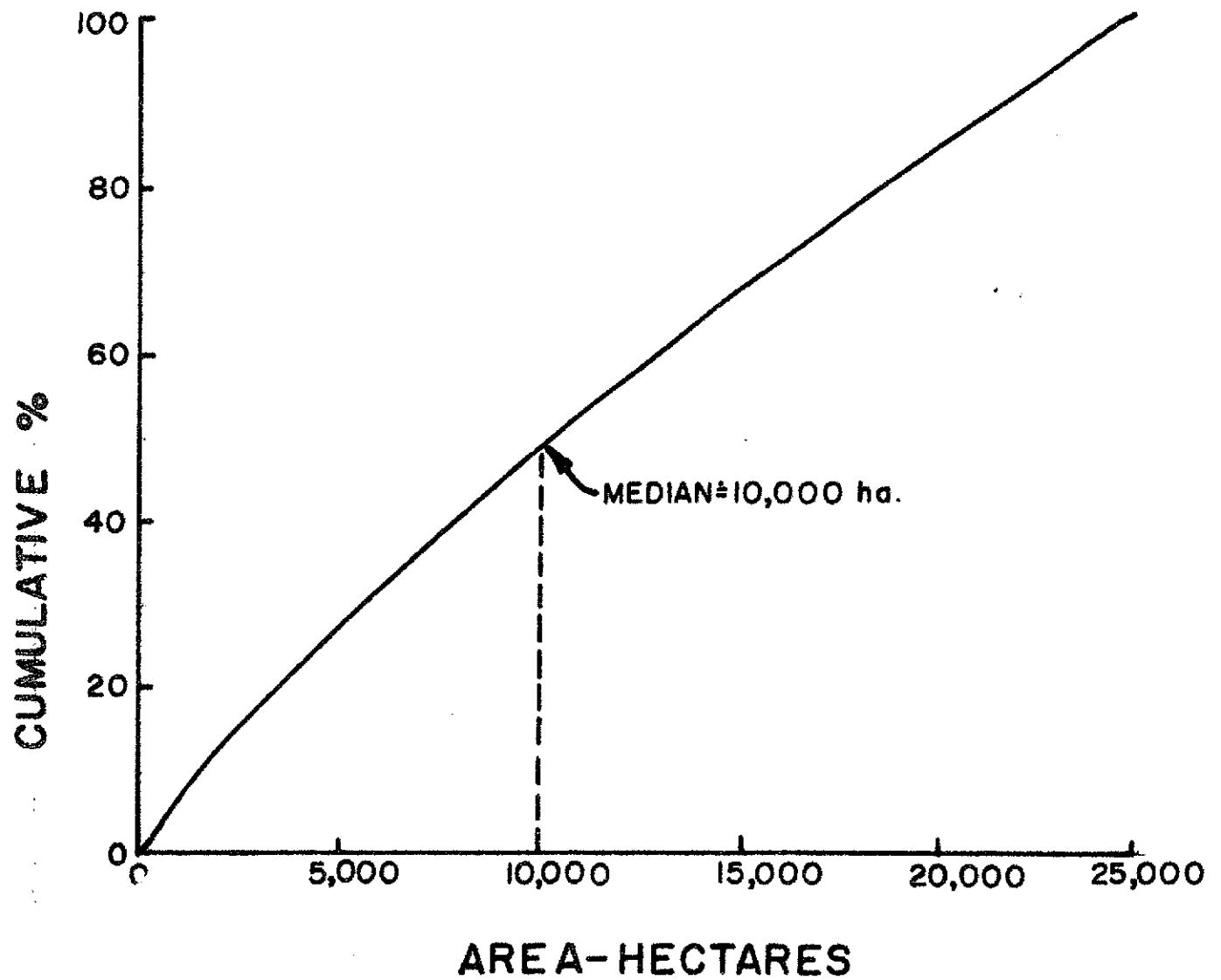
To illustrate: an add operation requires anywhere from three to five sequential elementary operations, known as stages. The exact number of stages depends upon the designer's option and the desired cost/performance. Each stage can be performed in a time commensurate with the switching time of the switching circuits. This time is approximately 10 nanoseconds for true, and tried low-cost technology; 3.5 nanoseconds for operational but costlier technology. Circuits can now be purchased, albeit at higher cost, with stage times as low as 2 nanoseconds. This means that a five-stage add can be performed currently in 50 nanoseconds with low-cost, 17.5 with medium-cost, and as low as 10 nanoseconds with high cost technology.

If, however, the program contains a string of adds, the second add can enter the multi-stage adder as soon as the first add has completed and cleared the first stage. This technique, known as pipelining, can cut the processing time down to the switching time of one stage.

Thus, for add operations a preprocessor can achieve speeds of 100 MIPS for low; 300 MIPS for medium, and 500 MIPS for high-cost technology. A multiply requires from five to ten stages:

FIGURE 10

WATERSHED AREA DISTRIBUTION- LOCAL AND STATE USERS



a divide, up to 30. Either can use the pipelining technique. It is clear that the average speed of a preprocessor will be a function of the "information entropy" of the program: the greater the number of elementary operations that can be arranged in sequence and pipelined, the higher the effective speed. The preprocessor output is buffered and fed as a summary to the general processor, which only performs the "synthesis operations."

By this means, image analysis by a general-purpose computer can be speeded up.

It is obvious that the preprocessor is most effective when used in conjunction with the slower machines. For example, a 100-MIP preprocessor would do little good on a 100-MIP machine.

Typically, on a 1-MIP machine such as the 360/75, a state-of-the-art preprocessor can cut the image processing time by a factor of approximately 40, thus reducing the time to process one ERTS frame from 10 hours to 15 minutes for simple processing, 30 minutes for sophisticated processing.

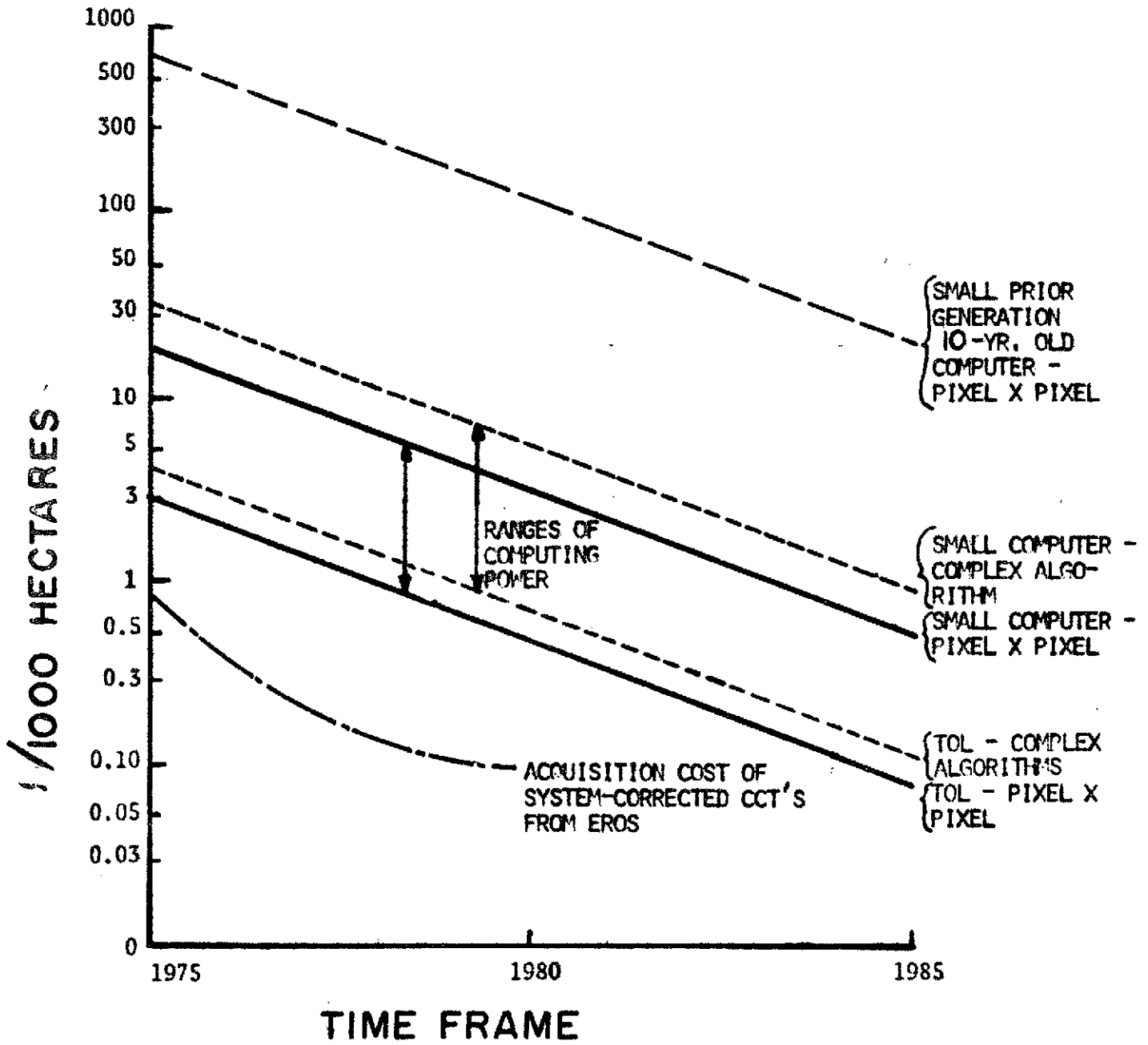
For very small machines, the preprocessor is also of limited velocity, because it has to "wait" for the machine to catch up after each batch of preprocessed instructions is fed to it.

The cost of preprocessing is expected to drop with time but not in step with the historical drop in data processing costs illustrated previously. The reason is that preprocessors are specialized devices, with far more limited market than general-purpose computers.

Figure 11 depicts the cost of processing ERTS computer-compatible tapes for hydrologic purposes, on general-purpose computers,

FIGURE 11

ERTS CCT-IMAGE PROCESSING COST/1000ha. WITHOUT PREPROCESSING



per 1,000 hectares of watershed, under the following two alternate conditions: 1) pixel by pixel classification, and 2) sophisticated processing.

Also indicated on Figure 11 are the acquisition costs of computer compatible tape (CCT) per 1,000 hectares. These costs are currently approximately \$225 per complete ERTS scene in four bands: they are expected to drop to \$100 by mid-1975, thence drop further with time, to an estimated \$50 per scene in the 1980 time frame. Note that at present CCT's are sold only on a per-scene basis.

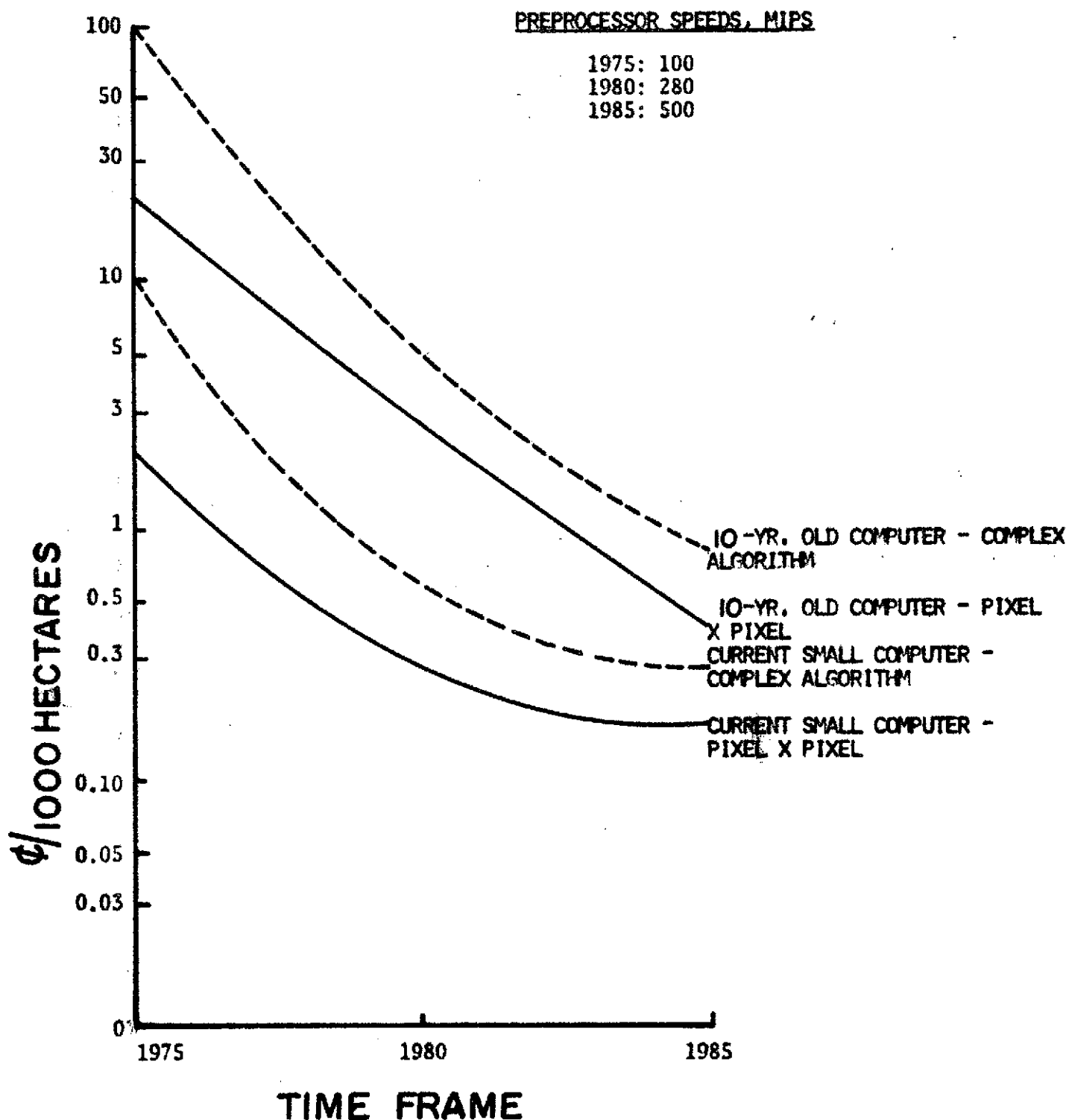
The cost trends shown in Figure 11 apply to "current" machines, i.e. computers of the latest models, whether large or small. Shown for comparison is also the cost situation for the smaller users, who utilize older machines. Note that the processing costs for the older machines are considerably higher, because their processing speeds are slow and the rental prices do not decrease in proportion to age. For example, the 360/30 which is now approaching 10 years of age since first entry to market, is still used rather widely for hydrologic modeling by small users.

Figure 12 depicts the processing costs achievable by addition of a typical preprocessor. The assumption made is that current commercially-available preprocessors have speeds of 100 MIPS equivalent: those of 1980, 280 MIPS: those of 1985, 500 MIPS. Although faster preprocessors could be custom-made, the corresponding investment would only be warranted by a very large, continuous applications load.

The costs shown in Figure 12 apply to current small machines,

FIGURE 12

ERTS CCT-IMAGE PROCESSING COST/1000ha. WITH PREPROCESSOR



which follow the trend depicted in Figure 11, and 10-year old machines. The costs of adding preprocessors to TOL machines is not shown since no significant speed improvements and thus cost savings do result.

Figures 11 and 12 apply to the data stream from ERTS. It is highly likely that the post-ERTS remote sensing data will obey the historical law of expanding use (or, in more popular parlance, Parkinson's law).

We will here concern ourselves with the growth in the complexity and consequent processing costs of remotely sensed imagery. Microwave radiometry, synthetic aperture radar and other more advanced applications are not treated in this effort.

To a first approximation, the number of instructions required to classify a pixel is directly proportional to a number of grey levels, inversely proportional to the square of the geometric resolution, directly proportional to the square dimension of the total area scanned, and directly proportional to the number of spectral bands.

$$i \propto \frac{nl^2f}{d^2} \quad (1)$$

where:

i = number of instructions

l = linear dimension of area scanned

f = number of spectral bands

n = number of grey levels

d = linear dimension of pixel

There exists, however, a fundamental relationship between the number of grey levels and linear pixel dimension, with all other system parameters remaining constant:

$$\frac{d^4}{n^2} = \text{const.} \quad (2)$$

Combining the above two relationships (1) and (2):

$$i \propto \ell^2 f \quad (3)$$

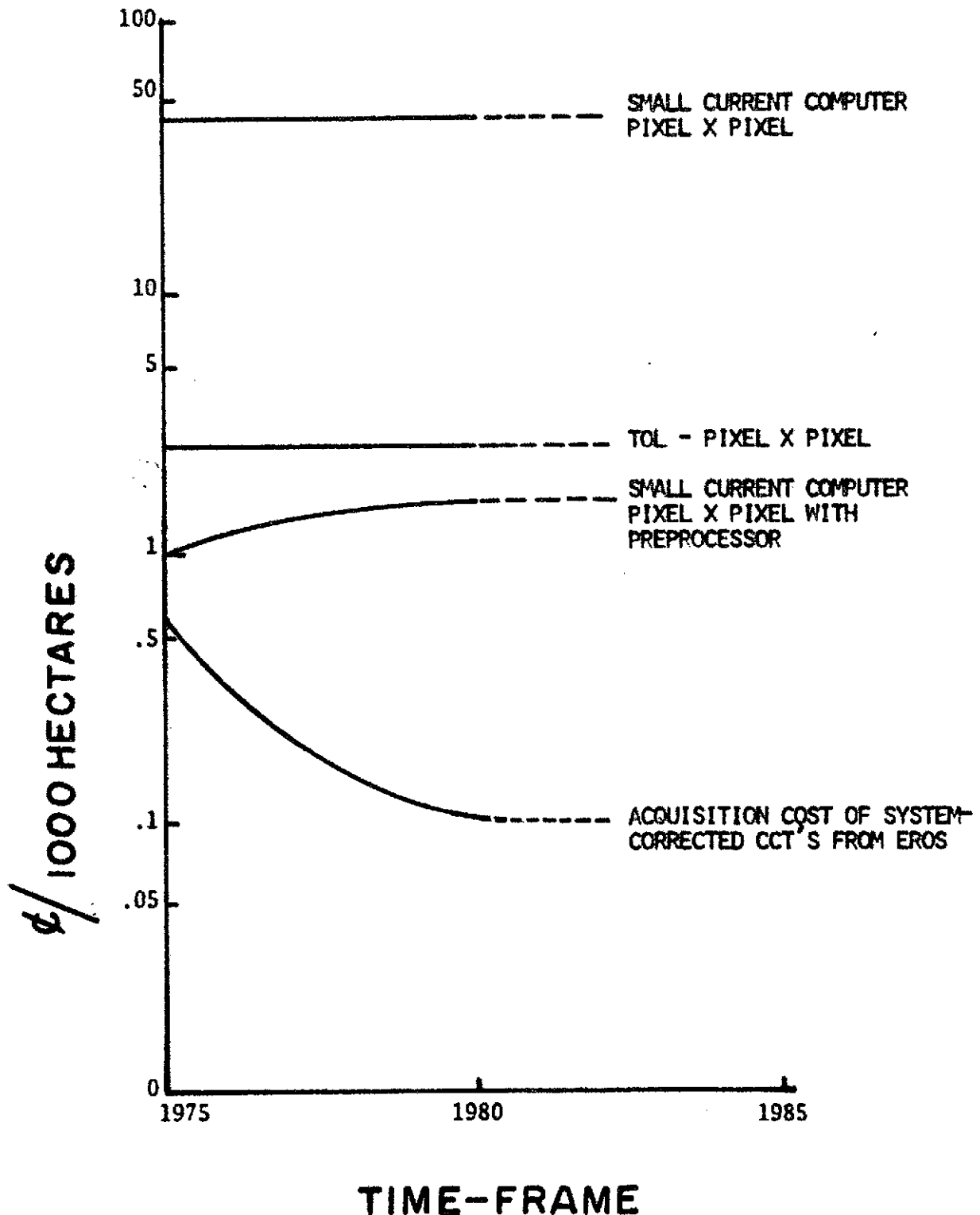
Note that equation (3) holds only for system parameters equal to those of ERTS: aperture size, orbital velocity and altitude, detector sensitivity, single sensor package.

Thus, a first step in the growth of data load will be caused by the addition of spectral bands: from the present 4 to the future 6: factor 1.5. Increases in detector sensitivity and aperture size combined, of approximately 12 db from the present MSS system can be reasonably anticipated by 1980. This is a further factor of 4. Thus, by approximately 1980, a total increase in data processing load of up to a factor of 6 for earth-orbiting remote sensors can be reasonably anticipated. Figure 13 depicts this trend.

Note that the CCT processing costs remain essentially constant. It is further interesting to note that the addition of a preprocessor to the smaller computers tends to increase the cost. The reason lies in the assumed growth pace of preprocessors, slower than the growth of general-purpose computing power.

FIGURE 13

GROWTH OF POST-ERTS REMOTELY
SENSED CCT IMAGERY PROCESSING
COST/1000ha.



Section 8

IMPACT OF THE REMOTE SENSING DATA STREAM UPON HYDROLOGIC USERS

From the material developed in the preceding sections, the following conclusions emerge:

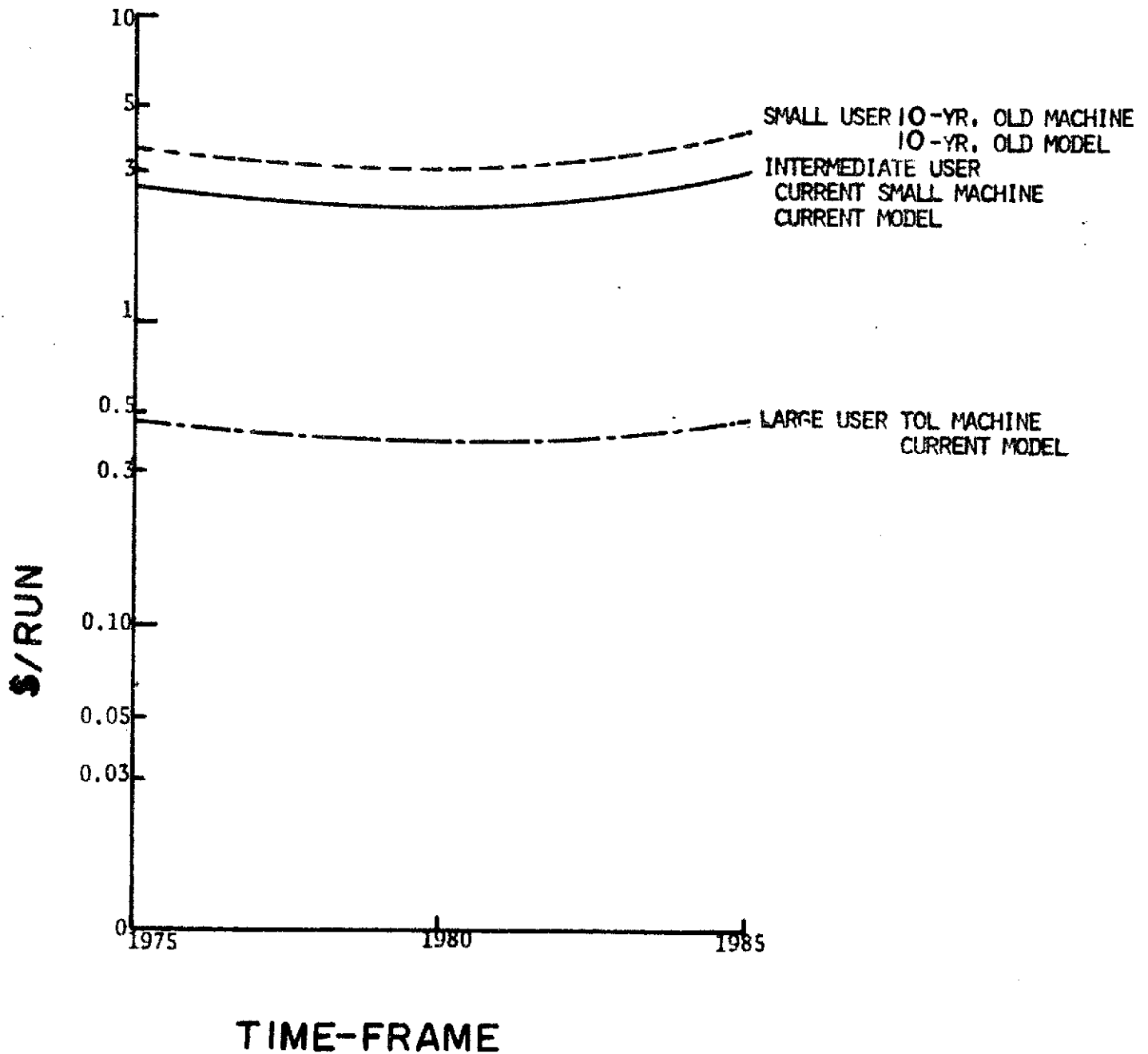
- 1) The cost of processing hydrologic models will remain substantially constant with time, as shown in Figure 14. This is a result of the contrasting trends of decreasing unit costs of data processing and increasing complexity of hydrologic models. The assumptions made in constructing Figure 14 are that the small user typically employs older, small machines, but also older-generation hydrologic models. The intermediate user employs current small machines and current, or almost-current, models. The larger user employs the best most powerful machines and the latest models.

It is clear that numerous variations can exist in machine-model combinations: the important characteristic is that they are contained within the region bounded by the upper and lower curves.

- 2) The cost of processing CCT's of the type currently produced by ERTS, in spite of substantial decreases, will be high. For the typical 10,000 hectare watershed, which represents the median of the small users, the processing cost is of order \$2.40 now. This is of the same order of magnitude as the cost of a run of his hydrologic model. If the small user continues to use the current type of ERTS CCT's, his processing cost by 1980 will drop to \$0.40.

FIGURE 14

**TREND OF DATA PROCESSING COSTS
FOR HYDROLOGIC MODEL RUN**



Thus, only by 1980 will the cost of CCT processing represent a reasonably low fraction of the cost of processing the model.

For the large user employing machines of power close to that of the TOL, but also dealing with much larger watersheds, say of order 1 million hectares, the CCT processing costs now would be of order \$40, much higher than the costs of processing the hydrologic model. Even in 1980 his costs would only drop to approximately \$6, still appreciably higher than model-processing costs.

- 3) The costs of procuring ERTS CCT's far outstrip the costs of CCT processing and of hydrologic modeling for all but the very large users.
- 4) The costs of processing CCT imagery for the remotely sensed data stream issuing from advanced post-ERTS earth-orbiting satellites will further accentuate the discrepancy between image-processing and model-processing costs.
- 5) As a consequence of the high costs of processing CCT's, the small and intermediate users will be significantly impacted in their effective use of remotely sensed imagery. Only the very large users with watersheds of areas approaching one ERTS frame will be able to take full advantage of the remotely sensed data.
- 6) The alternate potential consequences are threefold: 1) the small and intermediate users will significantly lag the large user in taking advantage of the remotely sensed data stream; or 2) they will resort to the more economical method

of direct analysis from imagery, suffering the consequent disadvantage of only partial utilization of the full gamut of information contained in the radiometric data; or 3) they will have to be served by some form of centralized facility, able to convert in the CCT's into information usable by the users. For example, supplying CCT data in mini-tape format encompassing this watershed alone. The magnitude and implications of these tradeoffs will be analyzed during the next reporting period.